

**DEVELOPMENT AND OPERATIONS OF THE  
ASTROPHYSICS DATA SYSTEM**

NASA Contract NCC5-528

Annual Report #3

For Period 1 January 2003 through 31 December 2003

Principal Investigator  
Dr. Stephen S. Murray

December 2003

Prepared for:

National Aeronautics and Space Administration  
Goddard Space flight Center  
Greenbelt, Maryland 20771

Smithsonian Institution  
Astrophysical Observatory  
Cambridge, Massachusetts 02138

The Smithsonian Astrophysical Observatory  
is a member of the  
Harvard-Smithsonian Center for Astrophysics

The NASA Technical Officer for this Grant is Ronald Oliverson, National Aeronautics and Space Administration, Goddard space flight Center, Greenbelt MD 20071

**SAO**

**ASTROPHYSICS DATA SYSTEM**

---

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 January 2003

**SUMMARY**

The ADS was represented at the AGU meeting with a poster paper and a demonstration. Discussions with the AGU about getting JGR abstracts continued.

Preparations for the ADS presence at the AAS meeting in January continued.

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 January 2003

## DEVELOPMENT

### SAO

#### TASKS ACCOMPLISHED:

##### Abstract Service

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Created new handlers for parsing abstracts/references for USNO, IBVS, and EGU journals.
- Created new handler and analyzed resolver results large IOP reference dataset (over 30 journals).
- Created correlation of publication date/volume for two large Elsevier physics journals (Nuclear Physics A/B), and processed table of contents for over 500 volumes.
- Implemented a prototype data verifier SOAP script as part of the ITWG activities on metadata exchange and verification.
- Optimized indexing procedures on new server to allow overnight indexing of most ADS databases.
- Created new parser for IOP references, downloaded back data and processed to recover references for entire 100,000 abstract dataset.
- Debugged various parsing issues with the IOP reference data set.
- Worked on mapping ISSN to bibstems for journals.
- Began to research available geophysics journals from major publishers to consider adding to a new Geophysics database.

##### Article Service

- Continued evaluation of commercial OCR software packages in preparation for full text OCR project.
- Created and sent 15 tapes containing article data to the Russian ADS mirror.

##### Historical Service

- Processed several complete series from the historical literature section.

##### Miscellaneous

- Responded to user comments about the Abstract and Article Services.
- Performed system administration/maintenance/repair on adsfife.
- Continued migration of software and documentation from tucana to adsfife.
- Worked on troubleshooting system problems with the China mirror.
- Debugged hardware and software problems with new Linux server, adsfife.

**SAO**

**ASTROPHYSICS DATA SYSTEM**

---

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 February 2003

**SUMMARY**

The ADS was represented at the AAS meeting with 3 poster papers and a demonstration booth.

We have setup a mirror site of the Vizier data base system at the CDS. This will replace the functionality of the ADC at Goddard.

Preparations for the APS and LPSC meetings in March started. We will have demonstrations at both meetings.

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 February 2003

## DEVELOPMENT

### SAO

#### TASKS ACCOMPLISHED:

##### Abstract Service

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Performed repairs and maintenance on various scripts which download and parse data automatically from the Web.
- Created new handlers for parsing html files downloaded from the Web for USNO, RSPTA, RSPSA, IBVS, and EGU journals.
- Created new handlers for parsing references from conferences and other smaller journals.
- Worked on improving the journal/journal abbreviation dictionary to improve reference parsing.
- Performed statistical analysis on quality of resolved references for physics journals from IOP and APS.
- Completed procedure allowing nightly updates of the ADS preprint database from the ArXiv holdings, including rapid indexing of over 200,000 records and group selection.
- Recreated all reference and citations after new injection of data from APS and IOP.
- Delivered ADS records for AAS journals to StreamSage, a company working in expert classification system as part of an ongoing collaboration.
- Wrote new parser for MmSAI data.
- Corresponded with IBVS personnel about various issues related to IBVS including refereed status, display of recent abstracts, and linking mechanism.
- Corresponded with PASA personnel to discuss issues about automating data delivery and to include references from old source files.
- Corresponded with IOP personnel and users about naming convention for the electronic-only JHEP journal.
- Provided SPIE personnel with abstracts from all records that we have of their journals.

##### Miscellaneous

- Responded to user comments about the Abstract and Article Services.
- Created XDOC parser for new XDOC format.
- Setup and documented US VizieR mirror on ADS machine.
- Performed system administration/maintenance/repair on various ADS machines.
- Continued debugging problems between hardware and software interaction on ADS's Linux server.
- Created handout for AAS meeting in Seattle, attended meeting to demonstrate new features of the ADS to users.

**SAO**

**ASTROPHYSICS DATA SYSTEM**

---

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 March 2003

**SUMMARY**

Preparations for the APS and LPSC meetings in March continued. We will have demonstrations at both meetings.

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 March 2003

## DEVELOPMENT

### SAO

#### TASKS ACCOMPLISHED:

##### Abstract Service

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Performed repairs and maintenance on various scripts which download and parse data automatically from the Web.
- Created scripts to automatically download PASA data from their website.
- Made improvements to APS resolver to improve reference resolution.
- Extracted and processed online references from all A&ARv articles available online at Springer.
- Created ISSN/journal mapping to be used for reference resolving on sources (e.g. IOP) where ISSN number is included in references.
- Created web-based snapshot overview of reference resolving in order to improve quality control.
- Fixed bugs in the preprint updating procedures.
- Identified ESASP conferences and created alternative bibcodes for them in preparation for scanning.

##### Article Service

- Processed and placed online over 100 scanned volumes of the ASP Conference Proceedings.
- Prepared and sent 9 boxes of publications to be digitized to our scanning contractor.

##### Miscellaneous

- Responded to user comments about the Abstract and Article Services.
- Performed system administration/maintenance/repair on various ADS machines.
- Migrated software from tucana to adsfife (reference resolver, mailing list for historical literature) to consolidate ADS services on fewer machines.
- Created snapshot overview tool for analysis of references.
- Continued to work on hardware and software issues related to Linux server.
- Continued maintenance of mirror sites with software and hardware problems, including mirrors in the UK, ESO, Brazil and China.
- Oversaw xeroxing and other computer tasks undertaken by summer student.

**SAO**

**ASTROPHYSICS DATA SYSTEM**

---

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 April 2003

**DISTRIBUTION**

(One copy for all recipients except as noted)  
Recipients please distribute within your group!

DISTRIBUTED BY SAO CONTRACTS:

NASA Hq/SZ:  
G. Riegler  
J. Hayes

DISTRIBUTED BY ADS PROJECT OFFICE:

PSU:  
J. Nousek

SAO/PROJ:  
S. Murray

### SUMMARY

The ADS was invited to organize a session about the ADS and its mirror sites at the next United Nations Workshop on Basic Space Sciences in the Developing World.

Efforts are under way to enter the tables of contents of all conference proceedings in the SAO library into the ADS. This requires copying the tables of contents from all volumes in the library and have them typed in. This will greatly enhance the coverage of the literature in the ADS.

We started the development of a search system for the full text of all scanned material in the ADS. This will eventually allow our users search capabilities that so far do not exist in any form.

In order to enable the full text searching, we have purchased OCR software and are in the process of OCRing the scanned pages in the ADS.

Efforts are in progress to handle the inclusion of data set identifiers in article manuscripts. The ADS will be the central system that will allow the journals to verify data set identifiers. The "master verifier" has been implemented in prototype form at the ADS.

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 April 2003

## DEVELOPMENT

### SAO

#### **TASKS ACCOMPLISHED:**

##### Abstract Service

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Performed repairs and maintenance on various scripts which download and parse data automatically from the Web, including rewriting IOP handler to use ISSN numbers, improving reference parsing for PThPh, repairing APS meetings downloading.
- Began systematic xeroxing of conference proceeding stacks in the CfA library to expand conference coverage in the astronomy database.
- Added almost 5000 abstracts from AGU, including all of JGR, GeoRL, and several other geophysics journals. This constituted all relevant records published in 2002.
- Added several hundred obituary entries to the ADS from USNO records provided to us.
- Added several thousand planetary records from Antarctic Meteorite Conferences provided to us by MetBase.
- Created listing of almost 1000 bibcode abbreviations to request from ISI in order to get physics citations, then pared that list down to about 600 of the most important journals.
- Added almost 1000 abstracts from Earth, Planets, and Science (EP&S).
- Gathered, checked and sent batches of tables of contents to send to India for typing. Special attention given to completing SPIE and ESA-SP conference proceedings tables of contents.
- Downloaded and parsed 150 volumes of CoPhC from Elsevier.
- Extracted and parsed references from A&ARv pdf files.
- Improvements to resolving for IAU Conferences, journal/journal abbreviation dictionary, and ISSN number dictionary.
- Updated configuration and data holdings for the ADS mirror sites in China, Brazil, and Russia after local hardware and software problems.
- Exchanged data with SPIRES to create cross-links between ADS and the SPIRES bibliographic archive.
- Updated procedure used to exchange references and citation data with the American Physical Society.

##### Data Verification Service

- Researched interoperability of SOAP toolkits for implementing data verification services among the NASA Astrophysics Data Centers.

##### Article Service

- Further development XDOC parser for new XDOC format.
- Supervision and troubleshooting of OCR process for full text.
- Prepared and shipped 9 boxes of journals to the scanning company

##### Miscellaneous

- Responded to user comments about the Abstract and Article Services.
- Oversaw xeroxing and other computer tasks undertaken by summer student.
- Created handout for LPI meeting in Houston, attended meeting to demonstrate new features of the ADS to users.
- Attended LPSC conference and gave demonstrations to participants.

- System administration/maintenance/repair for adsfife, including mail routing issue for Historical Lit. mailing list and access restriction to internal documentation.

**SAO**

**ASTROPHYSICS DATA SYSTEM**

---

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 May 2003

**DISTRIBUTION**

(One copy for all recipients except as noted)  
Recipients please distribute within your group!

DISTRIBUTED BY SAO CONTRACTS:

NASA Hq/SZ:  
G. Riegler  
J. Hayes

DISTRIBUTED BY ADS PROJECT OFFICE:

PSU:  
J. Nousek

SAO/PROJ:  
S. Murray

**SAO**

**ASTROPHYSICS DATA SYSTEM**

---

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 May 2003

**SUMMARY**

The ADS was represented with a poster at the joint AGU/EGU meeting in Nice, France.

Discussions about the on-going collaboration between the ADS and the CDS in Strasbourg, France were held in Strasbourg.

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 May 2003

## DEVELOPMENT

### SAO

#### **TASKS ACCOMPLISHED:**

##### Abstract Service

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Performed repairs and maintenance on various scripts which download and parse data automatically from the Web, including CoPhC, PhyS, and PASA.
- Built-in delay in HTTP requests of harvesting process since some servers deny access if request rate is too high.
- Created new downloading procedures for geophysics and other journals, including JApMe, JAtS, BAMS, JCli, JON, JOptT, LNP.
- Worked on parsing of USNO staff reference list.
- Improved reference parsing for A&A and A&AS, and improved resolving dictionaries.
- Continued systematic xeroxing of conference proceeding stacks in the CfA library to expand conference coverage in the astronomy database.
- Added over 11,000 abstracts from the last year of APS meetings.
- Analyzed AIP data which is lacking now that we no longer download from them but rely on STI to give us AIP data.
- Gathered, checked and sent 9 batches of tables of contents to send to India for typing. Special attention given to completing SPIE and ESA-SP conference proceedings tables of contents.

##### Data Verification Service

- Continued testing development of a SOAP PERL toolkit to allow interoperability of Dataset verification web services among NASA Astrophysics Data Centers
- Installed and tested PERL, Python, Java and PHP SOAP client/server software to test interoperability issues. Lead discussion on the issue within the ADEC ITWG (Interoperability Technical Working Group).

##### Article Service

- Supervision and troubleshooting of OCR process.
- Coordinated new schedule for microfilm digitization in collaboration with the SAO Wolbach Library and OCLC, the vendor providing digitization services for publication on microfilms.

##### Miscellaneous

- Responded to user comments about the Abstract and Article Services.
- System administration/maintenance/repair on adsfife, dealing with ".sit" under Unix, relocation of servers in computer room, configuration of new server "adswon".

**SAO**

**ASTROPHYSICS DATA SYSTEM**

---

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 June 2003

**DISTRIBUTION**

(One copy for all recipients except as noted)  
Recipients please distribute within your group!

DISTRIBUTED BY SAO CONTRACTS:

NASA Hq/SZ:  
G. Riegler  
J. Hayes

DISTRIBUTED BY ADS PROJECT OFFICE:

PSU:  
J. Nousek

SAO/PROJ:  
S. Murray

### SUMMARY

The ADS was invited to organize a session about the ADS and its mirror sites at the next United Nations Workshop on Basic Space Sciences in the Developing World.

Efforts are under way to enter the tables of contents of all conference proceedings in the SAO library into the ADS. This requires copying the tables of contents from all volumes in the library and have them typed in. This will greatly enhance the coverage of the literature in the ADS.

We started the development of a search system for the full text of all scanned material in the ADS. This will eventually allow our users search capabilities that so far do not exist in any form.

In order to enable the full text searching, we have purchased OCR software and are in the process of OCRing the scanned pages in the ADS.

Efforts are in progress to handle the inclusion of data set identifiers in article manuscripts. The ADS will be the central system that will allow the journals to verify data set identifiers. The "master verifier" has been implemented in prototype form at the ADS.

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 June 2003

## DEVELOPMENT

### SAO

#### **TASKS ACCOMPLISHED:**

##### Abstract Service

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Performed repairs and maintenance on various scripts which download and parse data automatically from the Web, including AsBio, ACP.
- Continued systematic xeroxing of conference proceeding stacks in the CfA library to expand conference coverage in the astronomy database.
- Performed a "citation adjustment", reparsing many of the citations in the ADS, analyzing what was no longer included, and responding to user questions about apparent loss of citations.
- Made additional improvements to the reference resolver, in particular with A&A and A&AS references. Configured and processed new conferences.
- Defined new category "6" for resolving references that can never be resolved such as "in press" and "private communication".
- Analyzed Geophysics data in the Physics database and elsewhere to assess whether to create a new Geophysics database for Planetary Scientists.
- Created mapping of MNRAS to their article ids in order to be able to create scans from recent PDFs.
- Added many missing ESA-SP conferences and created alternate bibcodes for thousands already included in the ADS.
- Gathered, checked and sent batches of tables of contents to send to India for typing. Special attention given to completing SPIE and ESA-SP conference proceedings tables of contents.

##### Data Verification Service

- Attended the International Virtual Observatory Alliance (IVOA) spring meeting in Cambridge, UK and gave a talk on the Verification of Data Identifiers via Web services, focusing on interoperability issues.
- Collaborated with University of Chicago Press and the AAS publication board in specifying the planned deployment of data identifiers in the astronomical literature.

##### Article Service

- Supervision and troubleshooting of OCR process.
- Installed OCR software and patches on adswon.
- Designed indexing and search architecture to be used for fulltext indexing of ADS's OCR sources.
- Prepared publications for scanning (Earth Moon and Planets, Celestial

##### Miscellaneous

- Responded to user comments about the Abstract and Article Services.
- Interviewed two candidates and hired one to work as a summer student.
- Wrote grant proposal to SI Seidell Grant program for hiring assistant to process historical literature scans.
- System administration/maintenance/repair (setup of RAID-5 on adswon).
- Upgraded new server to 4GB ram, tested performance of the PERL scripts used within ADS in new configuration.

**SAO**

**ASTROPHYSICS DATA SYSTEM**

---

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 July 2003

**DISTRIBUTION**

(One copy for all recipients except as noted)

Recipients please distribute within your group!

DISTRIBUTED BY SAO CONTRACTS:

NASA Hq/SZ:  
G. Riegler  
J. Hayes

DISTRIBUTED BY ADS PROJECT OFFICE:

PSU:  
J. Nousek

SAO/PROJ:  
S. Murray

### SUMMARY

We started to include more journals in Geosciences/Geophysics in the ADS.

The Royal astronomical Society has decided to archive their on-line journals in the ADS three years after publication. We have started to process these older on-line articles in order to archive them in the ADS.

Our mirror site in Korea now has a full article mirror.

We developed XML output capability in the ADS. This will make it easier to exchange data with other data systems.

We started the development of new indexing software that will eventually reduce the indexing time for a database from days to hours or less.

**DEVELOPMENT****SAO****TASKS ACCOMPLISHED:****Abstract Service**

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Performed repairs and maintenance on various scripts which download and parse data automatically from the Web, in particular geophysics journals including JApMe, BAMS, AnGla, JGla, AdG.
- Resolved more issues associated with the "citation adjustment".
- Continued systematic xeroxing of conference proceeding stacks in the CfA library to expand conference coverage in the astronomy database.
- Gathered, checked and sent batches of tables of contents to send to India for typing.
- Downloaded IoP data not previously included so that we can index their full set of journals.
- Corresponded with organizer and wrote paragraph to include in NASA Archives listings.
- Reworked creation of Blackwell links to work with their journals -- previously A&G links were broken.
- Created list of several thousand duplicate bibcodes from SIMBAD so that they can tell us which is correct and we can clean up outdated bibcodes from the database.
- Ported several scripts and modules to ADS's new linux servers in preparation of a transition to processing data under linux.

**Data Verification Service**

- Collaborated with the Chandra Archive to create and maintain data links to their mission and a Chandra group in ADS's search interface.
- Released a new version of the PERL toolkit used for dataset verification.

**Article Service**

- Set up process for reference extraction from XDOC (Solar Physics and Ap&SS).
- Integrated the archival and normalization of PDF and Postscript files supplied by publishers into ADS's article service.
- Met with staff at the contractor providing scanning services to the ADS project and discussed ways to improve their services through hardware upgrades and staffing changes.

**Miscellaneous**

- Responded to user comments about the Abstract and Article Services.
- System administration/maintenance/repair.
- Attended Special Libraries Association meeting and gave an ADS update at the Astronomy Roundtable. Met with librarians from other organizations and institutions.

**SAO**

**ASTROPHYSICS DATA SYSTEM**

---

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 August 2003

**DISTRIBUTION**

(One copy for all recipients except as noted)  
Recipients please distribute within your group!

DISTRIBUTED BY SAO CONTRACTS:

NASA Hq/SZ:  
G. Riegler  
J. Hayes

DISTRIBUTED BY ADS PROJECT OFFICE:

PSU:  
J. Nousek

SAO/PROJ:  
S. Murray

**SAO**

**ASTROPHYSICS DATA SYSTEM**

---

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 August 2003

**SUMMARY**

The ADS was represented at the IAU General Assembly with a poster. Discussions with the IAU management were held about extending the ADS IAU collaborations.

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 August 2003

## DEVELOPMENT

### SAO

#### TASKS ACCOMPLISHED:

##### Abstract Service

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Performed repairs and maintenance on various scripts which download and parse data automatically from the Web.
- Processed CDs with EGU abstracts.
- Maintenance of resolver scripts, including several new conferences and journals.
- Continued systematic xeroxing of conference proceeding stacks in the CfA library to expand conference coverage in the astronomy database.
- Gathered, checked and sent batches of tables of contents to send to India for typing.
- Corresponded with IoP personnel about inclusion of JCAP and JHEP, journals with non-standard volume/page pagination.
- Corresponded with PDS personnel about data to be included in the ADS, including naming schemes, and author vs. editor issues.
- Included approximately 3000 abstracts from the Australia IAU meeting.
- Corresponded with users, library personnel, and Springer personnel to solve the problem of the 4 year gap of electronic articles no longer available electronically.
- Updated and enhanced software used to match ADS and ISI's records in order to allow ingest of citation data from ISI into ADS's databases.

##### Article Service

- Reference extraction from XDOC (Solar Physics).
- Supervision and troubleshooting of OCR process
- Received, processed and placed online volumes 1-50 of "Astronomische Nachrichten" (AN).
- Wrote software to allow the creation and management of custom pagination for scanned articles. Trained ADS staff and summer student to use the software with scans for AN.
- Backed up and recovered data from the ADS article server after a faulty disk had to be replaced.

##### Miscellaneous

- Responded to user comments about the Abstract and Article Services.
- System administration/maintenance/repair (adsfife issues, tucana issues)
- Moved boxes of books to Metropolitan Storage which are not yet ready for scanning due to incompleteness or priorities.
- Investigated the web-crawling procedures used by google to index ADS's abstract data.

**SAO**

**ASTROPHYSICS DATA SYSTEM**

---

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 September 2003

**DISTRIBUTION**

(One copy for all recipients except as noted)  
Recipients please distribute within your group!

DISTRIBUTED BY SAO CONTRACTS:

NASA Hq/SZ:  
G. Riegler  
J. Hayes

DISTRIBUTED BY ADS PROJECT OFFICE:

PSU:  
J. Nousek

SAO/PROJ:  
S. Murray

**SAO**

**ASTROPHYSICS DATA SYSTEM**

---

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 September 2003

**SUMMARY**

The ADS staff moved offices from 60 Garden Street to the Porter Exchange building. There were no disruptions of the ADS services.

We started development of a new user interface that will allow weekly notification of new information in the ADS to users that request it.

Discussions with the PDS resumed about including more links from the ADS to the PDS.

Preparations for the demonstration at the DPS meeting in September started.

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 September 2003

## DEVELOPMENT

### SAO

#### **TASKS ACCOMPLISHED:**

##### Abstract Service

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Performed repairs and maintenance on various scripts which download and parse data automatically from the Web.
- Downloaded and parsed table of contents for SurSc from Elsevier.
- Improved reference resolving for AcASn and conferences.
- Gathered, checked and sent batches of tables of contents to send to India for typing.

##### Article Service

- Supervision and troubleshooting of OCR process (time-out problems during OCR).
- Definition of quality parameter for OCR results.
- Prepared publications for scanning, including Vilnius Observatory
- Publications, NInfo, Acta Astronomica, and Australian Journal of Physics.
- Moved over 50 boxes of journals and books to the warehouse for storage, created an inventory of all journal holdings for the project.
- Processed and made available online v. 1-274 of "Astronomische Nachrichten".
- Processed PDF files for MNRAS articles published in the years 1998-2001.

##### Miscellaneous

- Responded to user comments about the Abstract and Article Services.
- System administration/maintenance/repair (adsfife issues, jadzia crash)
- Move of all offices to PX and getting new work space ready.

**SAO**

**ASTROPHYSICS DATA SYSTEM**

---

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 October 2003

**DISTRIBUTION**

(One copy for all recipients except as noted)

Recipients please distribute within your group!

DISTRIBUTED BY SAO CONTRACTS:

NASA Hq/SZ:

G. Riegler

J. Hayes

DISTRIBUTED BY ADS PROJECT OFFICE:

PSU:

J. Nousek

SAO/PROJ:

S. Murray

**SAO**

**ASTROPHYSICS DATA SYSTEM**

---

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 October 2003

**SUMMARY**

The ADS was represented with a demonstration and a poster at the DPS meeting.

the master data set identifier verifier at ADS is working. Discussions with the ADEC and the AAS editorial office continued about the implementation of the tagging of data set identifiers in article manuscripts.

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 October 2003

## DEVELOPMENT

### SAO

#### **TASKS ACCOMPLISHED:**

##### Abstract Service

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Performed repairs and maintenance on various scripts which download and parse data automatically from the Web.
- Gathered, checked and sent batches of tables of contents to send to India for typing.
- Corresponded with librarians to straighten out some publication date questions in the ADS for AN, PALIO and others.
- Corresponded with data providers (EDP, Elsevier, Springer) to arrange for back issues of physics journals.
- Corresponded with data providers about data problems and linking issues(IoP), and new data formats (Elsevier).
- Corresponded with conference editors to establish our archiving and electronic linking policy.
- Debugged cookie-setting problem on Safari browsers.
- Wrote prototype C programs to be used as drop-in replacement for some of the indexing scripts suffering from performance problems.
- Updated synonym file used by the abstract service.

##### Data Verification Service

- Participated in email discussions concerning the definition of IVOA identifiers in order to provide consistency and compatibility with the ADEC ITWG implementation of dataset identifiers.

##### Article Service

- Performed reference extraction for SoPh and Ap&SS from OCR results.
- Harvested references from Level5.
- Investigated running resolver in update mode.
- Investigated access of OCR'd papers for historical literature site (MNRAS).

##### Miscellaneous

- Responded to user comments about the Abstract and Article Services.
- System administration/maintenance/repair (install of AMANDA backup server and clients).
- Attended Division of Planetary Sciences meeting and gave demonstrations at conference.

**SAO**

**ASTROPHYSICS DATA SYSTEM**

---

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 November 2003

**DISTRIBUTION**

(One copy for all recipients except as noted)  
Recipients please distribute within your group!

DISTRIBUTED BY SAO CONTRACTS:

NASA Hq/SZ:  
G. Riegler  
J. Hayes

DISTRIBUTED BY ADS PROJECT OFFICE:

PSU:  
J. Nousek

SAO/PROJ:  
S. Murray

**SAO**

**ASTROPHYSICS DATA SYSTEM**

---

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 November 2003

**SUMMARY**

Preparations for the new ADS notification service continued. This new service will be made available at the AAS meeting in January, 2004.

Preparations for the full text searching of the scanned material in the ADS continued. this service too will be available at the AAS meeting in January, 2004.

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 November 2003

## DEVELOPMENT

### SAO

#### **TASKS ACCOMPLISHED:**

##### Abstract Service

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Performed repairs and maintenance on various scripts which download and parse data automatically from the Web.
- Continued to try to extract EGU meeting abstracts from 2001/2002 CDROMs.
- Updated LaTeX macro definitions for reference resolver handlers to improve resolving.
- Collected data to create new group in the ADS for Leiden Observatory personnel.
- Gathered, checked and sent batches of tables of contents to send to India for typing.
- Worked on merging master lists from multiple databases into one master list for consistency across databases and to prepare for moving indexing onto adswon for a significant improvement in speed.
- Established a policy for determining whether or not a publication is refereed, in response to questions about refereed status of IAUC, Ap&SS, and others.
- Included approximately 13,000 new abstracts from EDP physics journals over the last decade or so.
- Helped offer suggestions to look and feel of new myADS service.
- Corresponded with Markus to improve journal matching and other resolver issues.
- Renamed approximately 15,000 AIPC abstracts to use AIPC as the primary bibcode.
- Worked on xeroxing physics books (QC) and journals in Wolbach library to improve coverage in physics database.
- Continued the porting and optimization of indexing software to linux servers.

##### Data Verification Service

- Prepared and presented the poster "Publishing Links to Astronomical Data On-line" at the ADASS XIII conference in Strasbourg, and participated to the IVOA workshop following the ADASS meeting.
- Updated the documentation about the Dataset Verification toolkits maintained by ADS.

##### Article Service

- Extracted abstracts for SoPh and Ap&SS from OCT results.
- Tested and analyzed impact of running resolver in update mode, including logging results.
- Implemented access of OCR'ed papers for historical literature site (MNRAS).
- Performed OCR on CeMec, CeMDA, Moon, AN.
- Further investigation in quality control for OCR results.

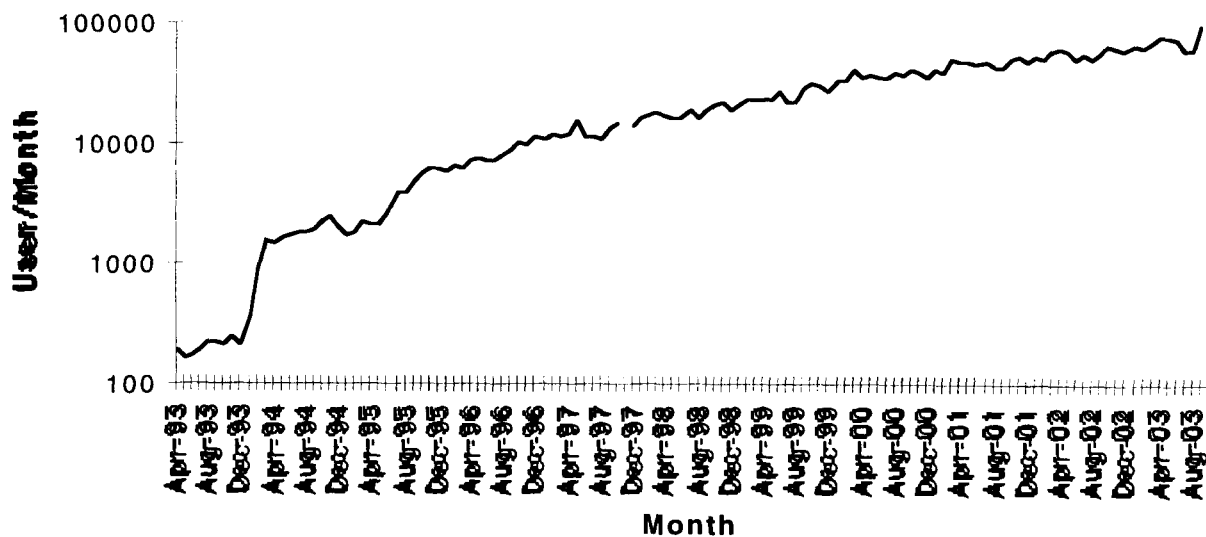
##### Miscellaneous

- Responded to user comments about the Abstract and Article Services.
- System administration/maintenance/repair (adsfife crashes, moved tucana to PX and attached scanner).
- Wrote abstract for AAS meeting.
- Reconfigured ADS servers in order to maintain our services available during the hardware failure of our main abstract server; developed procedures to allow easy recovery in case of additional hardware problems.

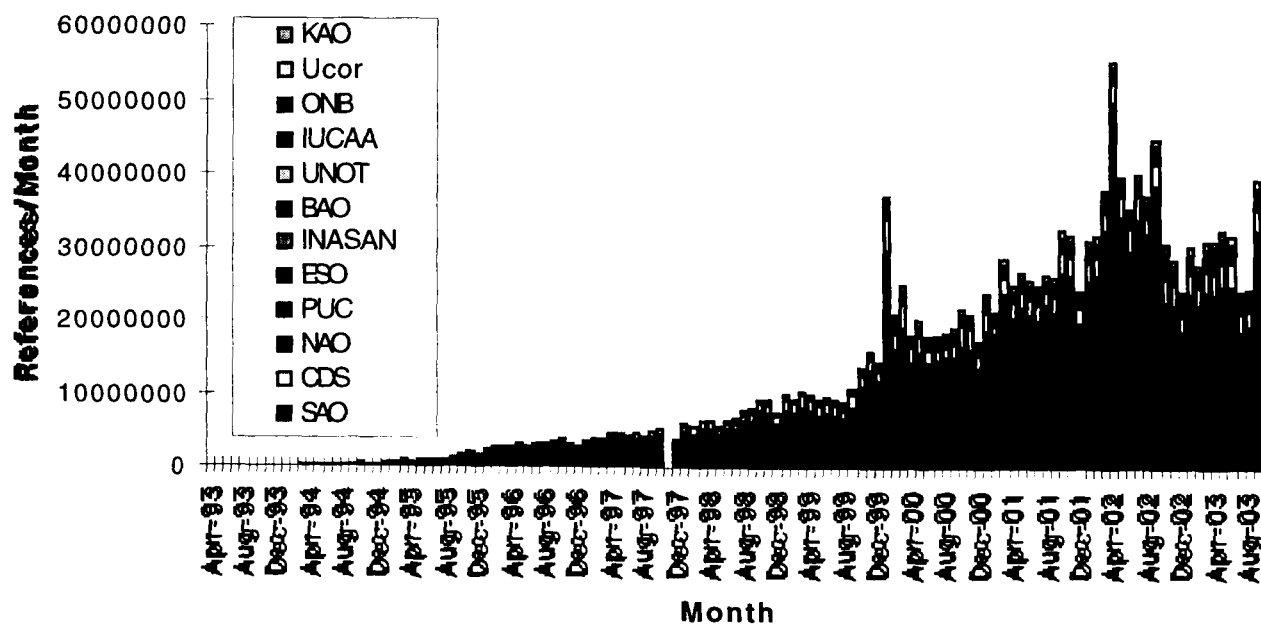
Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 December, 2003

### Distinct Users, Abstract Service



### Number of Retrieved References





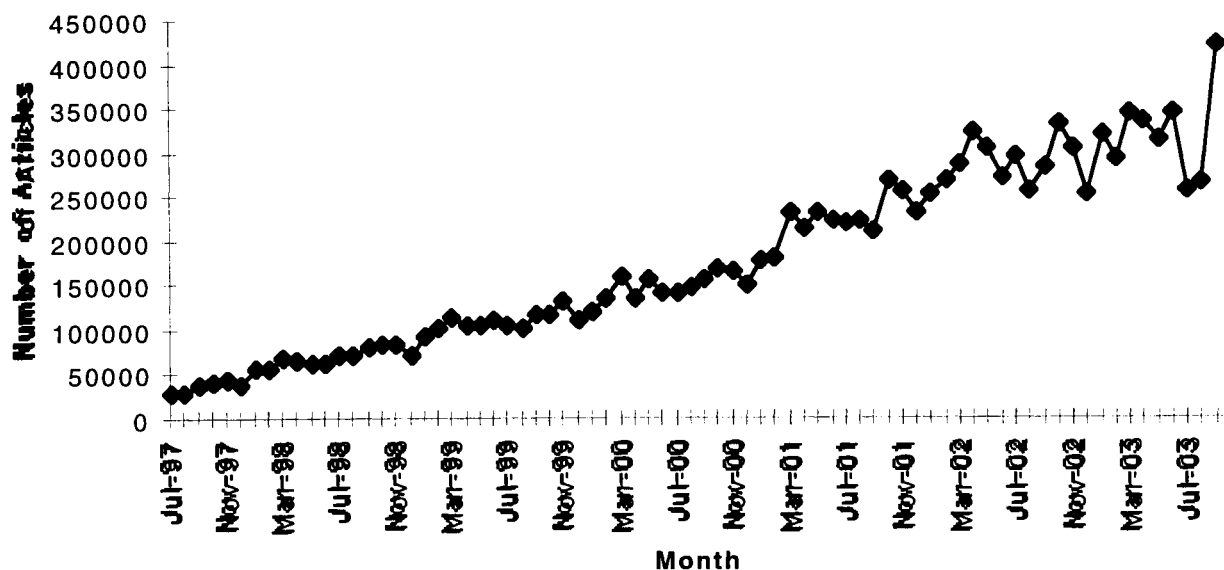
SAO

# ASTROPHYSICS DATA SYSTEM

Approved: \_\_\_\_\_ S. Murray  
Achievement: \_\_\_\_\_ G. Eichhorn (SAO)

Status as of: 1 December, 2003

## Number of Articles



## Number of Pages

